



The Bumps and BaBies Longitudinal Study (BaBBLeS): a multi-site cohort study of first-time mothers to evaluate the effectiveness of the Baby Buddy app

Deave, T., Ginja, S., Goodenough, T., Bailey, E., Piwek, L., Coad, J., Day, C., Nightingale, S., Kendall, S., & Lingam, R. (2019). The Bumps and BaBies Longitudinal Study (BaBBLeS): a multi-site cohort study of first-time mothers to evaluate the effectiveness of the Baby Buddy app. *mHealth*, 1-13.
<https://doi.org/10.21037/mhealth.2019.08.05>

[Link to publication record in Ulster University Research Portal](#)

Published in:
mHealth

Publication Status:
Published (in print/issue): 25/09/2019

DOI:
[10.21037/mhealth.2019.08.05](https://doi.org/10.21037/mhealth.2019.08.05)

Document Version
Publisher's PDF, also known as Version of record

General rights
Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.

The Bumps and BaBies Longitudinal Study (BaBBLeS): a multi-site cohort study of first-time mothers to evaluate the effectiveness of the Baby Buddy app

Toity Deave^{1#}, Samuel Ginja^{2#}, Trudy Goodenough¹, Elizabeth Bailey³, Lukasz Piwek⁴, Jane Coad³, Crispin Day⁵, Samantha Nightingale³, Sally Kendall⁶, Raghu Lingam⁷

¹Centre for Academic Child Health, Faculty of Health & Applied Sciences, University of the West of England Bristol, Bristol, UK; ²School of Psychology, Faculty of Life & Health Sciences, Ulster University, Coleraine, Northern Ireland; ³Centre for Innovative Research Across the Life-Course (CIRAL), Coventry University, Coventry, UK; ⁴Division of Information, Decisions and Operations, School of Management, University of Bath, Bath, UK; ⁵King's Health Partners, Child & Adolescent Mental Health Service Research Unit, Guy's Munro Centre, London, UK; ⁶Centre for Health Services Studies, University of Kent, Canterbury, UK; ⁷School of Women's & Children's Health, University of New South Wales, Randwick, New South Wales, Australia

Contributions: (I) Conception and design: T Deave, R Lingam, J Coad, S Kendall, C Day; (II) Administrative support: T Goodenough, S Ginja; (III) Provision of study materials or patients: T Deave, T Goodenough, S Ginja, E Bailey, S Nightingale; (IV) Collection and assembly of data: T Goodenough, S Ginja, L Piwek, E Bailey, S Nightingale; (V) Data analysis and interpretation: S Ginja, T Deave, E Bailey, S Nightingale, J Coad, R Lingam, L Piwek, C Day, S Kendall; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Toity Deave. iHV Fellow, Assoc. Professor for Family and Child Health, Centre for Academic Child Health, University of the West of England, Bristol, 1-5 Whiteladies Road, Bristol BS8 1NU, UK. Email: Toity.deave@uwe.ac.uk.

Background: Health mobile applications (apps) have become very popular, including apps specifically designed to support women during the ante- and post-natal periods. However, there is currently limited evidence for the effectiveness of such apps at improving pregnancy and parenting outcomes. This study aims to assess the effectiveness of a pregnancy and parenting app, Baby Buddy, in improving maternal self-efficacy at 3 months post-birth.

Methods: Participants were 16 years old or over, first-time mothers, 12–16 weeks gestation, recruited by midwives from five English study sites. The Tool to Measure Parenting Self-Efficacy (TOPSE) (primary outcome) was used to compare mothers at 3 months post-birth who had downloaded the Baby Buddy app with those who had not downloaded the app, controlling for confounding factors.

Results: Four hundred and eighty-eight participants provided valid data at baseline (12–16 weeks gestation), 296 participants provided valid data at 3 months post-birth, 114 (38.5%) of whom reported that they had used the Baby Buddy app. Baby Buddy app users were more likely to use pregnancy or parenting apps (80.7% *vs.* 69.6%, $P=0.035$), more likely to have been introduced to the app by a healthcare professional ($P=0.005$) and have a lower median score for perceived social support (81 *vs.* 83, $P=0.034$) than non-app users. The Baby Buddy app did not elicit a statistically significant change in TOPSE scores from baseline to 3 months post-birth [adjusted odds ratio (OR) 1.12, 95% confidence interval (CI): 0.59 to 2.13, $P=0.730$]. Finding out about the Baby Buddy app from a healthcare professional appeared to grant no additional benefit to app users compared to all other participants in terms of self-efficacy at 3 months post-birth (adjusted OR 1.16, 95% CI: 0.60 to 2.23, $P=0.666$). There were no statistically significant differences in the TOPSE scores for the in-app data, in terms of passive use of the app between high and low app users (adjusted OR 0.82, 95% CI: 0.21 to 3.12, $P=0.766$), nor in terms of active use (adjusted OR 0.47, 95% CI: 0.12 to 1.86, $P=0.283$).

Conclusions: This study is one of few, to date, that has investigated the effectiveness of a pregnancy and early parenthood app. No evidence for the effectiveness of the Baby Buddy app was found. New technologies

can enhance traditional healthcare services and empower users to take more control over their healthcare but app effectiveness needs to be assessed. Further work is needed to consider: (I) how we can best use this new technology to deliver better health outcomes for health service users and, (II) methodological issues of evaluating digital health interventions.

Keywords: Evaluation; first-time parents; Baby Buddy; self-efficacy; maternal well-being.

Received: 30 April 2019; Accepted: 08 August 2019; Published: 25 September 2019.

doi: 10.21037/mhealth.2019.08.05

View this article at: <http://dx.doi.org/10.21037/mhealth.2019.08.05>

Introduction

Electronic (e-Health) and mobile (m-Health) health methodologies are increasingly used to improve the self-management of health problems in many countries (1). This change in health seeking behaviour has been influenced by easier internet access, greater device functionality and poorer access to face-to-face healthcare services. There has been a growing interest in the capability of smartphone applications (“apps”) to promote health, encourage behaviour change and enhance the service users’ experience. There are over 318,000 health apps currently available on the leading app stores, with more than 200 apps added daily (2). However, systematic reviews have demonstrated that evidence of the effectiveness of health behaviour change apps remains limited and that studies of better quality are needed (3-5).

Ante- and post-natal care are two of the domains that have seen a large expansion of mobile apps. There are thousands of apps focused on women’s health and pregnancy, corresponding, approximately, to 7% of all existing health apps (6). It is commonly assumed that such apps have the potential to enhance conventional pregnancy and postnatal care (7). However, consistent with the wider literature on health apps, two systematic reviews found limited evidence of the effectiveness of apps designed specifically for ante- and/or post-natal care or women’s health (8,9). Although these reviews found a small number of evaluation studies where an experimental design had been used, they stressed the need for more high-quality studies and with adequately powered samples, as well as the need to assess the validity of app contents. It was also reported that, whilst some pregnancy and parenting app types have been assessed in a number of studies (e.g., gestational weight gain prevention), others, such as mental health-related apps, are lacking (9).

The Baby Buddy app was developed by the national child health and well-being charity, “Best Beginnings”. Its public health purpose was to provide evidence-based, professionally

validated information to pregnant and new mothers, empower women’s positive pregnancy and early parenting health behaviours, promote contacts with healthcare professionals and increase mothers’ self-efficacy with regard to pregnancy, baby care and early parenthood (10).

Parental well-being and self-efficacy, that is, parents’ self-perception about their ability to perform as parents, are major determinants of child health and development, parent-child relationships and buffer against parenting stress (11-13). The app content and functionality were co-created with parents and professionals and had a minimum reading age of 11 years with a “read aloud” element available. It included interactive information to help parents manage their physical and mental health and to help them to support the physical and emotional health of their child. It was designed to complement maternity and postnatal services and support the aim of “making every contact count” (14). Integration with health service delivery was promoted by Best Beginnings on the basis that mothers introduced to the app by a healthcare professional may be more likely to use it.

Based on “proportionate universalism” (15), Baby Buddy was intended to be used by mothers across the age-range with a particular focus on engaging groups at higher risk of poorer outcomes, such as expectant mothers under 25 years old. These younger mothers are less likely to engage with maternity services early in pregnancy and less likely to attend maternity appointments (16). Both behaviours are risk factors for adverse pregnancy outcomes (17). Baby Buddy was available for download by expectant mothers, partners, family members and friends from Apple iStore and the Google Play. Download data recorded by the app developers appeared to support its use by younger mothers (10).

The aim of the Bumps and BaBies Longitudinal Study (BaBBLes) reported in this paper was to assess the effectiveness of the Baby Buddy app on improving maternal self-efficacy and mental well-being at 3 months post-birth.

Methods

This longitudinal, mixed methods study was conducted in five geographical sites in England. It had three component parts: a cohort study, analysis of in-app data and a qualitative study. The study protocol has been previously published (18). An Appreciative Approach was used for the qualitative study with the results published elsewhere (19). The current paper reports on findings from the cohort study and in-app data analysis.

The cohort study compared self-reported, self-efficacy and mental well-being of (I) mothers 3 months post-birth who had used the Baby Buddy app with those mothers who had not, and (II) mothers who were shown how to use the app by a health professional, as suggested by the app developers, compared to those who were not shown or did not download it. In-app data were collected on uptake, usage pattern and detailed analytics of key app functionality.

Recruitment took place between September 2016 and February 2017. Women aged 16 years old and over, with no previous live child, and between 12–16 weeks and 6 days gestation were identified by the participating maternity units in the five study sites. Each identified woman was sent or given a study invitation letter and information booklet. Mothers completed questionnaires, online or on paper, which comprised of quantitative outcome measures and sociodemographic questions. A £5 voucher was issued upon receipt of the completed questionnaire (<http://fp.amegroups.cn/cms/mhealth.2019.08.05-1.pdf>). A 2-week reminder was sent if no questionnaire was received.

Data collection

Cohort study

Quantitative data were collected at three time points: 12–16-week pregnancy (baseline), 35-week pregnancy and 3 months post-birth. This paper focusses on the data collected at baseline and at 3 months post-birth. The inclusion of the 35-week gestation data did not affect these results significantly. All data were obtained from participant self-report.

At baseline, women provided informed consent for cohort study participation and completed the required measures.

In-app data

At the 35-week gestation data collection, mothers were sent an information sheet and consent form to complete in order to take part in this element of the study. The majority of Baby

Buddy app use patterns were recorded and stored on secured databases, hosted by Best Beginnings, as part of a standard procedure necessary for managing and debugging the app. For those mothers who gave their consent, using anonymised personal identification codes, Best Beginnings provided the research team with limited and secured download access to the database to obtain specific in-app data from app users, including duration of app use sessions, app session count, app use flow, and general user information.

Outcome measures

Primary outcome

Tool to Measure Parenting Self-Efficacy (TOPSE) (13,20).

The primary cohort study outcome measure was the TOPSE which is underpinned by self-efficacy theory (21). The TOPSE shorter version is a multi-dimensional instrument of 36 items within six scales representing distinct dimensions of parenting: emotion and affection, play and enjoyment, empathy and understanding, pressures, self-acceptance, learning and knowledge. The items are rated on an 11-point Likert scale, 0 (completely disagree) to 10 (completely agree), responses are summed to create a total score, lower scores indicating lower parenting self-efficacy. Subscale internal reliability coefficients ranged 0.80 to 0.89 and overall scale reliability was 0.94. External reliability coefficients ranged from $r_s = 0.58$ ($n=19$, $P<0.01$) to $r_s = 0.88$ ($n=19$, $P<0.01$). The 0–6-month version of TOPSE was adapted, in collaboration with the author, to measure parenting self-efficacy expectations during pregnancy.

Secondary outcome

Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) (22).

The WEMWBS was the secondary outcome measure validated for use in the UK with those aged 16 and above. It is a 14-item scale of subjective mental well-being and psychological functioning describing feelings (e.g., “I have been feeling useful”) and functional aspects (e.g., “I’ve been dealing with problems well”) over the previous 2 weeks. Items are scored from 1 (none of the time) to 5 (all of the time) and summed to provide an overall score between 14 and 70, where higher scores corresponded to greater mental well-being frequency. WEMWBS has good content and criterion-related validity and high test-retest reliability [0.83 (23)].

Sociodemographic variables

Sociodemographic and health data collected included

women's age, ethnic group, socio-economic deprivation, highest level of formal education, relationship status and employment. Index of multiple deprivation (IMD) decile, a common indicator of socioeconomic deprivation in the UK, was obtained by searching participants' postcodes using a standard online tool (24). The geographical site where participants were recruited was also noted. Social support was measured using the Multidimensional Scale of Perceived Social Support [MSPSS (25)] and technology use was assessed using the Media and Technology Usage and Attitudes Scale (MTUAS) (26). In addition, at baseline and at 35 weeks gestation, participants' expected date of delivery (EDD) and intended baby feeding methods were recorded. At 3 months post-birth, information about participants' childbirth experience, using the Childbirth Experience Questionnaire (CEQ) (27), and actual baby feeding methods were collected. For more details see the published protocol (18).

Sample size

Our original sample size calculation assumed linearity of outcome variables (18). Both primary and secondary outcomes were negatively skewed and therefore converted to dichotomous variables, lowest quartile compared to the upper three quartiles. The original sample size of 559 women assumed a 12.5% app download, which meant roughly a ratio of 1 Baby Buddy user to 7 non-users (18). However, as explained in the results section, the percent app download was higher than anticipated which reduced the required sample size to 250 participants (due to a smaller ratio). This included 100 intervention subjects (i.e., Baby Buddy app users) and 150 controls (i.e., non-app users) to have 80% power to detect a 7% difference [0.5 standard deviation (SD)] in the proportion of participants in the lowest quartile compared to the upper three quartiles at the 5% level (28).

Data analysis

Descriptive statistics were used to describe the sample, including the mothers' age, socio-demographics, ethnicity, access to and use of technology and the overall sum scores for the outcome measures. Logistic regression models were used to compare the primary and secondary outcomes in mothers who used the Baby Buddy app compared to those who did not use the app. Participants were considered app users if they had reported using the app at any of the three

data collection time points. Logistic regression diagnostics using Hosmer and Lemeshow's goodness-of-fit test indicated a good fit of the adjusted models ($P > 0.05$). Key variables were tested as potential confounders, including maternal age, education, employment, relationship status, recruitment site, social support, general technology use and use of other pregnancy apps. Baseline levels of the outcome variables were also controlled for in the final analysis. Analysis was as per protocol and analysis plan unless otherwise specified. All analyses were carried out using Stata 14 software (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP).

The TOPSE scores were negatively skewed so a log transformation of these data was carried out but the distribution remained non-normal. As a result, we developed logistic regression models in which TOPSE scores were converted into a binary variable: low self-efficacy [1], to represent those in the lowest quartile of TOPSE score data and reference levels of self-efficacy [0], which corresponded with those with TOPSE scores above the lowest quartile. In this analysis, we report the odds ratio (OR) of low TOPSE scores (i.e., low self-efficacy) amongst Baby Buddy app users compared to non-app users. This logistic regression analysis comprised of two models: (I) unadjusted model and, (II) model adjusted for potential confounders, including baseline levels of the outcome.

Secondary analysis

A second analysis compared primary and secondary outcomes, as described above, between those mothers who used the app and heard about it from a health professional (instructed use) and those women who did not hear about it or who did not download the app by 3 months post-birth.

Post-hoc analysis

Qualitative findings suggested that Baby Buddy breastfeeding contents were popular (19). It was decided to conduct a *post-hoc* analysis of the impact of the Baby Buddy app on self-reported breastfeeding.

In-app data

For consenting mothers ($n=51$), uptake, patterns of usage and detailed analytics of key factors within the app were analysed. These were participants who had provided valid outcome data at baseline (i.e., TOPSE or WEMWBS data) and who also responded at 3 months post-birth with valid outcome data.

Data orientation was undertaken and then formatted

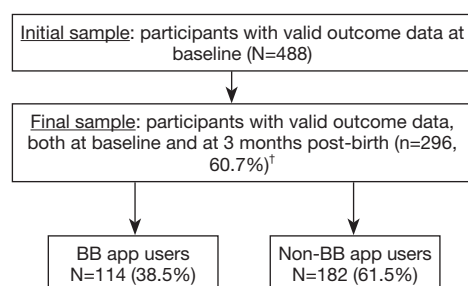


Figure 1 Participant flow in the BaBBLeS study. BaBBLeS, the Bumps and BaBies Longitudinal Study; BB, Baby Buddy. [†], 192 participants did not have valid outcome data at both baseline and 3 months post-birth.

for analysis. This included an exploratory analysis of socio-demographic information and profiling of app users (e.g., age, occupation, education, ethnic origin); description of app use patterns including the creation of the app avatar; goal setting function, media downloaded and the app functions of “ask me a question” and “what does that mean”.

In consultation with the app developers, the following app elements were assessed to quantify in-app usage: “Today’s Information”, “Videos”, “Ask Me”, “Remember to Ask”, “You can Do it”, “Bump Around/Baby Around”, “Baby Book/Bump Book”, “Baby Booth/Bump Booth”, and “What Does it Mean”. The number of times each element of the app was used were summed and two overall aggregated scores were derived for data analysis. The first score was a “passive” overall score, based exclusively on the “Today’s Information” element. This included whether this feature had been opened, if links were followed and whether participants tapped on “Read more”. This involved mostly viewing and clicking information and was less goal- and behaviour change-oriented. The second composite score was an “active” overall score and encompassed all other app elements. This was a more proactive format of app interaction, for example, users had to search specifically for information or videos or set up reminders.

Based on the median value of the session count, the passive users were sub-divided into passive high users (n=26; 94 sessions or more) and passive low app users (n=25; less than 94 sessions). Similarly, the active high app users (n=27; 27 sessions or more) and active low app users (n=24; less than 27 sessions) sub-divided into two groups. Separate logistic regression models were developed to compare outcomes (TOPSE and WEMWBS, as dichotomised in previous models) between active high and low app users and

passive high and low app users. The same two regression models used for the questionnaire data were performed, one unadjusted (model 1) and one adjusted for potential confounders (model 2). However, considering the small number of participants in the analyses, to maximise the viability of the model, there had to be careful selection of the confounding variables to be included. Differences between high/low app users were analysed and confounding factors were selected which were shown to be significant at the baseline outcome level for TOPSE and WEMWBS.

Ethics

This study received a favourable opinion from the NHS Research Ethics Committee (NRES) West Midlands-South Birmingham REC (16/WM/0029), the University of the West of England, Bristol Research Ethics Committee (HAS.16).

Results

Descriptive results

A total of 488 participants provided valid data at baseline, i.e., TOPSE data and/or WEMWBS data (initial sample). Of this initial sample, 256 participants (52.5%) provided valid data at 35 weeks gestation. Of the initial sample, 296 (60.7%) provided valid data at 3 months post-birth; this was the sample used in the main analysis, hereinafter referred to as the final sample. There were 220 participants (45.1%) who provided data at all three data collection time-points. The participant flow is presented in *Figure 1*.

Of the 296 participants followed to 3 months post-birth, 114 reported to be Baby Buddy app users (38.5%), i.e., they had reported using the Baby Buddy app at one or more of the three data collection time-points. This corresponds roughly to a ratio of 1 to 2, i.e., one reported Baby Buddy user for every two non-Baby Buddy users.

The distribution of participants in the initial sample (n=488) by recruitment site was as follows: 168 from the West Midlands (34.4%), 139 from London (28.5%), 66 from West Yorkshire (13.5%), 62 from Lancashire (12.7%) and 53 from East Midlands (10.9%). This distribution, per site, remained very similar in the final sample. Baseline characteristics of participants included in the final sample are presented by app use in *Table 1*. App users (n=114) were comparable to non-app users (n=182) in age, IMD decile, ethnicity, highest education attained, employment and relationship status.

Table 1 Baseline characteristics of participants by Baby Buddy use

Variable	Baby Buddy users ^a (n=114)			Non-Baby Buddy users (n=182)			Differences between app users and non-app users
	n missing	n (%)	Med (LQ–UQ)	n missing	n (%)	Med (LQ–UQ)	
Age (continuous)	4		29.5 [26–33]	3		29 [25–33]	Z=−0.85, P=0.396
16–24 years		17 (15.5)			41 (22.9)		
25–34 years		77 (70.0)			115 (64.2)		
35 years or above		16 (14.5)			23 (12.8)		
IMD decile (continuous)	1		4 [2–6]	4		4 [3–7]	Z=1.36, P=0.176
1 to 3 (most deprived)		49 (43.4)			67 (37.6)		
4 to 6		42 (37.2)			64 (36.0)		
7 to 10 (least deprived)		22 (19.5)			47 (26.4)		
Ethnicity	3		–	5		–	$\chi^2(1) = 3.23$, P=0.072
White British		67 (60.4)			125 (70.6)		
Other ^b		44 (39.6)			52 (29.4)		
Highest education	1		–	3		–	$\chi^2(1) = 0.08$, P=0.774
Degree or higher		65 (57.5)			106 (59.2)		
No degree		48 (42.5)			73 (40.8)		
Employment	3		–	1		–	$\chi^2(1) = 0.50$, P=0.479
In paid employment		97 (87.4)			163 (90.1)		
Not in paid employment		14 (12.6)			18 (9.9)		
Relationship	1		–	0		–	$\chi^2(1) = 0.23$, P=0.628
Married or living with partner		102 (90.3)			161 (88.5)		
Not married or not living with partner		11 (9.7)			21 (11.5)		

^a, Baby Buddy users are those who reported using the app at one or more data collection time-points; ^b, Ethnicity: “Other” includes White Irish (n=3), White Other European (n=32), White Other (n=9), Asian or Asian British Pakistani (n=17), Asian or Asian British Bangladeshi (n=3), Asian or Asian British Indian (n=13), Asian or Asian British Chinese (n=4), Asian or Asian British Other (n=6), Black or Black British Caribbean (n=3), Black or Black British African (n=4), and Mixed (n=2). Z, based on Mann-Whitney test; χ^2 , based on chi-squared test; n missing, number of missing responses; Med, median; LQ–UQ, lower quartile–upper quartile; IMD, index of multiple deprivation, based on postcode;

All participants used a mobile phone and had internet access and nearly all had internet at home. Two thirds used a tablet. There were no significant baseline differences between Baby Buddy users and non-Baby Buddy users in terms of any of these variables. The three top sources of information about pregnancy and parenthood, in both groups, were the internet (app users 88.5%; non-app users 82.7%), friends (app users 82.4%; non-app users 76.5%) and midwife (app users 74.3%; non-app users 71.0%). For both Baby Buddy users and non-Baby Buddy users, the overall median MTUAS score was 5. No

significant differences with regards to any of these variables were observed between the two groups. There were no differences in terms of use of technology scores between Baby Buddy users and non-app users (27).

Baby Buddy users were significantly more likely to use pregnancy/parenthood apps in general, not just the Baby Buddy app, than non-Baby Buddy users at baseline (80.7% vs. 69.6%, P=0.035) consequently, this was one of the variables adjusted for in the main analysis. Baby Buddy users were also more likely to have heard about the pregnancy apps they used from healthcare professionals than non-

Table 2 Baseline scores of TOPSE and WEMWBS

Variable	Baby Buddy users (n=114)		Non-Baby Buddy users (n=182)		Difference tests between app users and non-app users ^a
	n missing	Median [LQ–UQ]	n missing	Median [LQ–UQ]	
TOPSE subscale ^b					–
Emotion & affection	3	55 [50–58]	6	56 [52–59]	
Play & enjoyment	2	59 [55–60]	4	59 [56–60]	
Empathy & understanding	2	55 [48–58]	4	55 [50–60]	
Pressures	4	42 [33–48]	2	40 [32–50]	
Self-acceptance	2	54 [48–59]	6	55 [48–59]	
Learning & knowledge	2	53 [49–58]	1	56 [51–60]	
TOPSE overall	6	317 [287–337]	13	320 [295–337]	Z=1.03, P=0.302
WEMWBS overall ^c	2	54 [49–59]	3	54 [48–61]	Z=0.45, P=0.655

^a, based on Mann-Whitney test; ^b, TOPSE: Tool to Measure Parenting Self-Efficacy; 36 items, 6 subscales (6 items per subscale). Subscale score range is 0–60; Overall score range is 0–360. Items 6, 19, 20, 21, 27 were reverse scored. Higher scores (higher self-efficacy). An adapted antenatal version of the TOPSE was used at baseline and 35–40 weeks; the (already existing) postnatal TOPSE was used at follow-up. ^c, WEMWBS: Warwick-Edinburgh Mental Well-being Scale. Only one overall score 14–70. Higher scores = higher well-being. LQ, lower quartile; UQ, upper quartile.

Baby Buddy users (42.4% *vs.* 24.4%, $P=0.005$). On the overall MSPSS score, Baby Buddy users had a significantly lower median score than non-Baby Buddy users (81 *vs.* 83, $P=0.034$); this indicates lower levels of perceived social support amongst Baby Buddy users at baseline.

Baseline data for the outcome variables show that the median score for the TOPSE was 317 (LQ–UQ, 287–337) for app users 320 (LQ–UQ, 295–337) for non-app users (Table 2). For the WEMWBS, the median for app users and non-app users were 54 (LQ–UQ, 49–59) and 54 (LQ–UQ, 48–61), respectively. There were no statistically significant differences between the two groups for either the TOPSE or WEMWBS. Similar to the MSPSS, TOPSE and WEMWBS scores are used for comparison between participants or across time.

Outcome results

At 3 months post-birth, there were no statistically significant differences in TOPSE or WEMWBS outcomes between Baby Buddy users and non-Baby Buddy users. Baby Buddy users had a median TOPSE score of 319 [lower quartile (LQ)–upper quartile (UQ), 296–338] compared to non-Baby Buddy users who had a median TOPSE score of 327 (LQ–UQ, 305–343) ($P=0.107$). Similarly, Baby Buddy

users had a median WEMWBS score of 54.5 (LQ–UQ, 49–59) compared to non-Baby Buddy users who had a median score of 55 (LQ–UQ, 50–61) ($P=0.284$).

The unadjusted OR for low TOPSE score (i.e., lower self-efficacy) was 1.17 (95% CI: 0.68 to 2.03, $P=0.564$) amongst Baby Buddy users compared to non-Baby Buddy users (Table 3). Adjustment of this association for IMD decile, technology use (baseline MTUAS total mean score), use of pregnancy/parenthood apps (any), social support (baseline MSPSS overall sum score) and baseline TOPSE score resulted in a very similar result: adjusted OR of 1.12 (95% CI: 0.59 to 2.13, $P=0.730$). The Baby Buddy app had no significant effect on maternal mental well-being, with an unadjusted OR for low WEMWBS of 1.10 (95% CI: 0.64 to 1.89, $P=0.719$). Adjustment for confounding factors made minimal difference to this association, OR 1.02 (95% CI: 0.55 to 1.89, $P=0.943$) (Table 3).

Baby Buddy users who had heard about the app from a healthcare professional had slightly higher odds of a low self-efficacy TOPSE score compared to all other participants. These differences were not statistically significant, neither in the unadjusted model (model 1) (OR 1.16, 95% CI: 0.66 to 2.04, $P=0.596$) nor in the adjusted model (model 2) (OR 1.16, 95% CI: 0.60 to 2.23, $P=0.666$). Similarly, there were no differences in the ORs for low

Table 3 Odds ratios for low TOPSE scores and reported Baby Buddy use

Outcome	Model	N	Baby Buddy use		
			OR (SE)	95% CI	P value
TOPSE	Model 1	282	1.17 (0.33)	0.68 to 2.03	0.564
	Model 2	263	1.12 (0.37)	0.59 to 2.13	0.730
WEMWBS	Model 1	294	1.10 (0.30)	0.64 to 1.89	0.719
	Model 2	283	1.02 (0.32)	0.55 to 1.89	0.943

Model 1: Baby Buddy use (at any time) and outcome overall sum score at 3 months post-birth, unadjusted; Model 2: same as model 1, adjusted for IMD decile, technology use (baseline MTUAS total mean score), use of pregnancy/parenthood apps (any), social support (baseline MSPSS overall sum score), baseline outcome overall sum score. IMD, index of multiple deprivation; OR, odds ratio; SE, standard error; TOPSE, Tool of Parenting Self-efficacy; WEMWBS, Warwick-Edinburgh Mental Well-Being Scale; MTUAS, Media and Technology Usage and Attitudes Scale; MSPSS, Multidimensional Scale of Perceived Social Support.

Table 4 Number of uses of the app

Level of in-app usage	Passive use: overall score, n (%)	Active use: overall score, n (%)
Less than 5 uses	3 (5.9)	4 (7.8)
5–15 uses	4 (7.8)	9 (17.6)
16–25 uses	4 (7.8)	9 (17.6)
26–45 uses	7 (13.7)	13 (25.5)
46–100 uses	9 (17.6)	11 (21.6)
101–250 uses	17 (33.3)	5 (9.8)
251+ uses	7 (13.7)	0 (0)
Median [LQ–UQ]	94 [31–196]	27 [13–51]

LQ, lower quartile; UQ, upper quartile.

WEMWBS scores between Baby Buddy users who had heard about the app from a healthcare professional and all other participants, neither in the unadjusted model (OR 1.03, 95% CI: 0.59 to 1.79, $P=0.924$) nor in the adjusted model (OR 1.00, 95% CI: 0.53 to 1.87, $P=0.990$).

In-app data

The number of uses of each aggregated score: passive and active usage (*Table 4*), suggest that participants engaged more with the passive elements of the app.

Changes in levels of app usage and whether they affected the reported outcomes (i.e., TOPSE and WEMWBS scores) were explored. The differences between the characteristics of in-app participants (those who had consented to their in-app data being used and who had provided valid outcome data at baseline and 3 months post-birth ($n=51$) and non-Baby Buddy users ($n=182$) were similar to those differences

between Baby Buddy users and non-Baby Buddy users, i.e., statistically non-significant except that in-app users had lower social support ($P=0.035$) and used more pregnancy/parenthood apps than non-Baby Buddy users ($P<0.0001$).

The results of the logistic regression analysis for both self-efficacy (TOPSE) and mental well-being (WEMWBS) and any association with usage of the passive and active in-app elements are described in *Table 5*. For clarity, we also report the median value of the outcome score, for each of the two groups (under the columns “High users” and “Low users”). The results revealed no statistically significant associations between level of usage of the passive in-app element and TOPSE scores, and WEMWBS scores, neither in the unadjusted nor in the adjusted models. Confidence intervals were large, particularly for WEMWBS. Another set of analyses were performed comparing high app users with non-Baby Buddy users, rather than with low users. Results, not reported here, were very similar to

Table 5 In-app use and outcome data

In-app aggregated score	Outcome	High users (n=26)	Low users (n=25)	Model	OR (SE)	95% CI	P value
Passive use of the app	TOPSE overall	315 [302–330]	312 [296–329]	Model 1	0.81 (0.52)	0.23 to 2.88	0.747
				Model 2	0.82 (0.56)	0.21 to 3.12	0.766
	WEMWBS overall	54 [48–57]	55 [52–58]	Model 1	1.67 (1.05)	0.49 to 5.69	0.408
				Model 2	3.58 (2.77)	0.78 to 16.3	0.099
Active use of the app	TOPSE overall	319 [305–336]	309 [287–318]	Model 1	0.48 (0.31)	0.13 to 1.73	0.261
				Model 2	0.47 (0.33)	0.12 to 1.86	0.283
	WEMWBS overall	54 [48–57]	54.5 (49.5–58.5)	Model 1	1.50 (0.94)	0.44 to 5.09	0.516
				Model 2	3.50 (2.78)	0.74 to 16.5	0.112

The values under the “High users” and “Low users” columns correspond to median (lower quartile – upper quartile). Model 1 is unadjusted; Model 2 is adjusted for education and outcome scores at baseline. Outcomes are binary; odds of low TOPSE, odds of low WEMWBS, based on the overall median value, as used in previous analyses. WEMWBS, Warwick-Edinburgh Mental Well-being scale; TOPSE, Tool of Parenting Self-efficacy; OR, odds ratio; SE, standard error.

Table 6 Reported breastfeeding* in the final sample

Reported breastfeeding	Baby Buddy user (N=114), n (%)	Non-Baby Buddy user (N=182), n (%)	Differences between app users and non-app users
Any breastfeeding at 1 week post-birth	100 (87.7)	141 (79.2)	$\chi^2(1) = 3.49$, $P = 0.062$
Any breastfeeding at 1 month post-birth	95 (84.8)	121 (67.6)	$\chi^2(1) = 10.68$, $P = 0.001^*$
Any breastfeeding at 3 months post-birth	69 (61.6)	93 (51.4)	$\chi^2(1) = 2.93$, $P = 0.087$
Exclusive breastfeeding at 1 week post-birth	65 (57.0)	91 (51.1)	$\chi^2(1) = 0.97$, $P = 0.325$
Exclusive breastfeeding at 1 month post-birth	55 (49.1)	67 (37.4)	$\chi^2(1) = 3.86$, $P = 0.050$
Exclusive breastfeeding at 3 months post-birth	49 (43.8)	61 (33.7)	$\chi^2(1) = 2.98$, $P = 0.084$

*, breastfeeding includes both breastfeeding as the exclusive feeding method as well as breastfeeding in combination with formula milk, compared to formula milk only or not sure.

those presented in *Table 5*, with no statistically significant differences between the two groups.

Post-hoc analysis on breastfeeding

Baby Buddy users were more likely to report that they had breastfed at 1-week post-birth, at 1-month post-birth and at 3 months post-birth (*Table 6*). This included breastfeeding in combination with formula milk (“any breastfeeding”) and breastfeeding as the sole baby feeding method (“exclusive breastfeeding”). At 1-month post-birth, this difference was statistically significant for any breastfeeding [$\chi^2(1) = 10.68$, $P = 0.001$] (*Table 6*).

Logistic regression models were developed to explore the association between breastfeeding and Baby Buddy use, using the same unadjusted and adjusted models from

the main analysis (*Table 7*). At all time-points, Baby Buddy app users had increased odds of reported breastfeeding compared to non-Baby Buddy users. However, differences between the two groups were only statistically significant for any breastfeeding at 1 month post-birth, both unadjusted (OR 2.68, 95% CI: 1.46 to 4.90, $P = 0.001$) and after adjusting for confounding variables (OR 3.08, 95% CI: 1.49 to 6.35, $P = 0.002$) and at 3 months post-birth in the adjusted model for exclusive breastfeeding (OR 1.79, 95% CI: 1.02 to 3.16, $P = 0.044$) (*Table 7*).

Discussion

There is a lack of evidence about the effectiveness of pregnancy/parenthood apps, with those studies that

Table 7 Odds ratios for breastfeeding and Baby Buddy use

Reported breastfeeding	Model	N	Baby Buddy use		
			OR (SE)	95% CI	P value
Any breastfeeding at 1 week post-birth	Model 1	292	1.87 (0.64)	0.96 to 3.65	0.065
	Model 2	280	2.25 (0.93)	1.00 to 5.06	0.051
Any breastfeeding at 1 month post-birth	Model 1	291	2.68 (0.82)	1.46 to 4.90	0.001*
	Model 2	275	3.08 (1.14)	1.49 to 6.35	0.002*
Any breastfeeding at 3 months post-birth	Model 1	293	1.52 (0.37)	0.94 to 2.45	0.088
	Model 2	276	1.72 (0.49)	0.99 to 2.99	0.054
Exclusive breastfeeding at 1 week post-birth	Model 1	292	1.27 (0.31)	0.79 to 2.04	0.325
	Model 2	280	1.13 (0.30)	0.67 to 1.90	0.649
Exclusive breastfeeding at 1 month post-birth	Model 1	291	1.61 (0.39)	1.00 to 2.60	0.050
	Model 2	275	1.65 (0.45)	0.97 to 2.80	0.067
Exclusive breastfeeding at 3 months post-birth	Model 1	293	1.53 (0.38)	0.94 to 2.48	0.085
	Model 2	276	1.79 (0.52)	1.02 to 3.16	0.044*

Baby Buddy users are those who reported using the Baby Buddy app at one time-point or more. Any breastfeeding includes both breastfeeding as the exclusive feeding method as well as breastfeeding in combination with formula milk; exclusive breastfeeding refers to those women who reported breastfeeding as the only feeding method. Model 1: Breastfeeding and Baby Buddy use, unadjusted; Model 2: same as model 1, adjusted for IMD decile, education, technology use (MTUAS total mean score), use of pregnancy/parenthood apps (any), and baseline intention to breastfeed. *, $P < 0.05$. IMD, index of multiple deprivation; MTUAS, Media and Technology Usage and Attitudes Scale; OR, odds ratio; SE, standard error.

aimed to assess this being insufficiently powered to detect significant effects (8,9). The BaBBLLeS study aimed to address this research gap by being one of the first large-scale controlled studies to assess the effectiveness of such an app, Baby Buddy, at improving reported maternal psychological outcomes. Our findings suggested that the app had no effect on maternal parenting self-efficacy and mental well-being at 3 months post-birth. There were also no statistically significant outcome differences between those who used the app more than the median number of app sessions and those who used it less, based on objective (in-app) data, or between those who were told about the app by a healthcare professional and those who found out about it through other sources.

Although the use of the Baby Buddy app did not impact on the pre-specified outcomes, a *post-hoc* analysis suggested that it did lead to higher levels of self-reported breastfeeding, after adjusting for baseline differences and other relevant confounders. These findings, though preliminary, are hypothesis generating and potentially encouraging. Nevertheless, as a *post-hoc* analysis the findings require further exploration using a pre-specified plan of

analysis, ideally in a randomised controlled trial. This is particularly important given its relevance to the current public health agenda. The exploration of which specific features of the app are responsible for the improvements in breastfeeding would be helpful for healthcare practitioners, especially midwives and health visitors, so that those features could be emphasised in their contact with mothers.

Midwives were the most frequent source of information about Baby Buddy, suggesting that the app developers were successful in their maternity dissemination methods with the aim to “make every contact count” (29). However, findings suggested that the app may not lead to the expected improvements in maternal self-efficacy and mental well-being even when integrated into in service delivery; improvements in non-hypothesised outcomes such as breastfeeding were detected.

The lack of expected outcome impact may be due to the absence of the interpersonal and personalised aspects of care that are core elements of face-to-face clinical interactions [e.g., (30,31)]. It may be that apps may have a supplementary role but are unlikely to replace direct clinical care especially when managing the challenges affecting

the lives of vulnerable women during pregnancy and early infancy (32,33).

Strengths and limitations of the study

Outcome data were based on self-report using well-validated scales used previously to detect significant increases in self-efficacy and mental well-being. The TOPSE was adapted for antenatal use and the effect of anticipated, compared to actual, self-efficacy, on post-birth optimism is unknown. Outcome scores on both TOPSE and WEMWBS were high at baseline in both app user group and the non-app user group, raising the potential of ceiling effects. There was little change in total scores at each time point, inferring that the participant cohort was generally high functioning in parenting self-efficacy and mental well-being. While the app may have sought to influence these outcomes, participants expressed preference for talking to healthcare professionals face-to-face and to be with other parents (19).

The study used a broad definition of “Baby Buddy user” that included any use of the app during the study period. This definition is consistent with an intention to treat approach but may lack sensitivity to the use of specific app functionality. The secondary analysis using the in-app data, however found no differences between high and low/no app users. This suggests that the lack of association between outcomes and Baby Buddy use was unlikely to have been due to measurement errors.

A longer, e.g., 6-month, follow-up period may have been preferable. However, a systematic review of web-based interventions for perinatal mood disorders suggests that 3-month follow-up assessments can detect outcome improvement (34).

Using a randomised, rather than quasi-experimental, design would strengthen the inferences drawn from the study’s findings. However, randomisation was not possible because the Baby Buddy app was freely available for download, risking contamination in those randomised to a comparison condition. Furthermore, one of the few differences between Baby Buddy app using and non-app using mothers at baseline was the use of other maternity apps by the Baby Buddy app-using mothers, which suggests that mothers may either be users of several apps or none (35).

We are unable to provide an estimate of the proportion of women approached by midwives who agreed to study participation. While using recruitment logs, maternity staff limitations, prevented them from being anonymised and then shared with the research team. Retention rates in

studies involving ante- and post-natal women are variable but the study’s 60% rate is consistent with those reported in clinical research trials involving perinatal women (36,37). It attests to the difficulty of engaging with new mothers at such a demanding period of their lives. The final sample included just those mothers who had complete data for the TOPSE and WEMWBS at baseline and at 3 months post-birth. The baseline characteristics of those mothers in the final sample largely reflected those of the initial sample and app users and non-app users remained comparable.

Participants were self-selected and we were unable to assess their representativeness for the wider population of first-time mothers in each site. The sample was predominantly composed of White British women living in areas of higher economic deprivation (38). However, the rate of degree holders, at baseline, 51.0% and in the final sample, 58.6%, is substantially higher than the national average of 42% (39). This was affected by the characteristics of the London site, where a considerable part of our sample was based. The greater likelihood of more socially advantaged participants is a common phenomenon in maternal health-related research (40,41).

Conclusions

There is an increasing emphasis on the use of technologies to support the delivery of healthcare services, as evident from the National Health Service apps library (42). New technologies may have potential to enhance and even replace conventional healthcare provision as well as empower people to take more control over their healthcare. This is one of the few studies to date to investigate the health outcomes of a specific app designed for use by mothers in the antenatal and early postnatal periods. It found no evidence of impact on first-time mothers’ self-reported parental self-efficacy and mental well-being at 3 months post-birth though *post-hoc* analysis suggested that app users were more likely to report to exclusively breastfeed, or ever breastfeed. Overall findings suggest that this particular app may have limited impact on the outcomes measured. Further work is needed to differentiate the types of outcomes the app may improve as well as how new technologies more widely can be best optimised to health outcomes.

Acknowledgments

The authors would like to thank all the participants of this

study—the mothers and the health professionals. They would also like to thank the five participating midwifery services who supported and undertook the process of recruitment to the study and follow-up data collection. This work was undertaken while the authors S Ginja and R Lingam were still based at the Institute of Health and Society, at Newcastle University. This work was supported by the Big Lottery via Best Beginnings as a competitive tender.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study received a favourable opinion from the NHS Research Ethics Committee (NRES) West Midlands-South Birmingham REC (16/WM/0029), the University of the West of England, Bristol Research Ethics Committee (HAS.16).

References

1. Zapata BC, Fernández-Alemán JL, Idri A, et al. Empirical studies on usability of mHealth apps: a systematic literature review. *J Med Syst* 2015. doi: 10.1007/s10916-014-0182-2.
2. IQVIA. The Growing Value of Digital Health [Internet]. 2017 [cited 2019 Mar 13]. Available online: <https://www.iqvia.com/institute/reports/the-growing-value-of-digital-health>
3. Zhao J, Freeman B, Li M. Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. *J Med Internet Res* 2016;18:e287.
4. Marcolino MS, Oliveira JAQ, D'Agostino M, et al. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. *JMIR Mhealth Uhealth* 2018;6:e23.
5. McKay FH, Cheng C, Wright A, et al. Evaluating mobile phone applications for health behaviour change: A systematic review. *J Telemed Telecare* 2018;24:22-30.
6. Aitken M, Lyle J. Patient adoption of mHealth: use, evidence and remaining barriers to mainstream acceptance. Parsippany, NJ: IMS Institute for Healthcare Informatics, 2015. Available online: <https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/patient-adoption-of-mhealth.pdf> (Accessed: 5 July 2019).
7. Tripp N, Hainey K, Liu A, et al. An emerging model of maternity care: Smartphone, midwife, doctor? *Women Birth* 2014;27:64-7.
8. Overdijkink SB, Velu AV, Rosman AN, et al. The Usability and Effectiveness of Mobile Health Technology-Based Lifestyle and Medical Intervention Apps Supporting Health Care During Pregnancy: Systematic Review. *JMIR Mhealth Uhealth* 2018;6:e109.
9. Derbyshire E, Dancey D. Smartphone Medical Applications for Women's Health: What Is the Evidence-Base and Feedback? *Int J Telemed Appl* 2013;2013:782074.
10. Best Beginnings. Best Beginnings. About Baby Buddy [Internet]. 2017 [cited 2018 Sep 17]. Available online: <https://www.bestbeginnings.org.uk/about-baby-buddy>
11. Coleman PK, Karraker KH. Maternal self-efficacy beliefs, competence in parenting, and toddlers' behavior and developmental status. *Infant Ment Health J* 2003;24:126-48.
12. Deave T, Heron J, Evans J, et al. The impact of maternal depression in pregnancy on early child development. *BJOG* 2008;115:1043-51.
13. Kendall S, Bloomfield L. Developing and validating a tool to measure parenting self-efficacy. *J Adv Nurs* 2005;51:174-81.
14. Percival J. Promoting health: making every contact count. *Nursing Standard* 2014;28:37-41.
15. Marmot MG, Allen J, Goldblatt P, et al. Fair society, healthy lives: Strategic review of health inequalities in England post-2010 [Internet]. London, UK: The Marmot Review, 2010 Feb [cited 2019 Apr 30]. Available online: <http://discovery.ucl.ac.uk/111743/>
16. Bradshaw P, Schofield L, Maynard L. The Experiences of Mothers Aged Under 20: Analysis of Data from the Growing Up in Scotland Study. Scottish Government Social Research, 2014.
17. Raatikainen K, Heiskanen N, Heinonen S. Under-attending free antenatal care is associated with adverse pregnancy outcomes. *BMC Public Health* 2007;7:268.
18. Deave T, Kendal S, Lingam R, et al. A study to evaluate the effectiveness of Best Beginnings' Baby Buddy phone app in England: a protocol paper. *Prim Health Care Res Dev* 2019;20:e19.
19. Deave T, Coad J, Day C, et al. Bumps and Babies Longitudinal Study (BABBLIES): An independent evaluation of the Baby Buddy app [Internet]. 2018 [cited 2019 Apr 30]. Available online: <http://eprints.uwe.ac.uk/39012/>
20. Bloomfield L, Kendall S. Parenting self-efficacy, parenting stress and child behaviour before and after a parenting

- programme. *Prim Health Care Res Dev* 2012;13:364-72.
21. Bandura A. Self-efficacy mechanism in human agency. *Am Psychol* 1982;37:122-47.
 22. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes* 2007;5:63.
 23. Stewart-Brown SL, Platt S, Tennant A, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a valid and reliable tool for measuring mental well-being in diverse populations and projects. *J Epidemiol Community Health* 2011;65:A1-40.
 24. English indices of deprivation 2015 [Internet]. GOV.UK. [cited 2019 Apr 30]. Available online: <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015>
 25. Zimet GD, Dahlem NW, Zimet SG, et al. The Multidimensional Scale of Perceived Social Support. *J Pers Assess* 1988;52:30-41.
 26. Rosen LD, Whaling K, Carrier LM, et al. The Media and Technology Usage and Attitudes Scale: An empirical investigation. *Comput Human Behav* 2013;29:2501-11.
 27. Dencker A, Taft C, Bergqvist L, et al. Childbirth experience questionnaire (CEQ): development and evaluation of a multidimensional instrument. *BMC Pregnancy Childbirth* 2010;10:81.
 28. Dupont WD, Plummer WD. Power and sample size calculations: A review and computer program. *Control Clin Trials* 1990;11:116-28.
 29. Best Beginnings. Enhancing Capacity of Professionals & Community [Internet]. 2017 [cited 2019 Mar 13]. Available online: <https://www.bestbeginnings.org.uk/enhancing-capacity-of-professionalscommunity>
 30. Seward MW, Simon D, Richardson M, et al. Supporting healthful lifestyles during pregnancy: a health coach intervention pilot study. *BMC Pregnancy Childbirth* 2018;18:375.
 31. Willcox JC, van der Pligt P, Ball K, et al. Views of Women and Health Professionals on mHealth Lifestyle Interventions in Pregnancy: A Qualitative Investigation. *JMIR MHealth UHealth* 2015;3:e99.
 32. Santarossa S, Kane D, Senn CY, et al. Exploring the Role of In-Person Components for Online Health Behavior Change Interventions: Can a Digital Person-to-Person Component Suffice? *J Med Internet Res* 2018;20:e144.
 33. Prentice JL, Dobson KS. A review of the risks and benefits associated with mobile phone applications for psychological interventions. *Can Psychol* 2014;55:282-90.
 34. Lee EW, Denison FC, Hor K, et al. Web-based interventions for prevention and treatment of perinatal mood disorders: a systematic review. *BMC Pregnancy Childbirth* 2016;16:38.
 35. Lupton D, Pedersen S. An Australian survey of women's use of pregnancy and parenting apps. *Women Birth* 2016;29:368-75.
 36. Frew PM, Saint-Victor DS, Isaacs MB, et al. Recruitment and retention of pregnant women into clinical research trials: an overview of challenges, facilitators, and best practices. *Clin Infect Dis* 2014;59 Suppl 7:S400-7.
 37. McCarter DE, Demidenko E, Hegel MT. Measuring outcomes of digital technology-assisted nursing postpartum: A randomized controlled trial. *J Adv Nurs* 2018;74:2207-17.
 38. Ginja S, Coad J, Bailey E, et al. Associations between social support, mental wellbeing, self-efficacy and technology use in first-time antenatal women: data from the BaBBLLeS cohort study. *BMC Pregnancy Childbirth* 2018;18:441.
 39. Office for National Services. Graduates in the UK labour market - Office for National Statistics, ONS. GOV.UK. Available online: <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/graduatesintheuklabourmarket/2017> (Accessed: 25 July 2019).
 40. Braig S, Grabher F, Ntomchukwu C, et al. The Association of Hair Cortisol with Self-Reported Chronic Psychosocial Stress and Symptoms of Anxiety and Depression in Women Shortly after Delivery. *Paediatr Perinat Epidemiol* 2016;30:97-104.
 41. Feinberg ME, Jones DE, Roettger ME, et al. Preventive Effects on Birth Outcomes: Buffering Impact of Maternal Stress, Depression, and Anxiety. *Matern Child Health J* 2016;20:56-65.
 42. NHS. NHS Apps Library [Internet]. 2019 [cited 2019 Mar 13]. Available online: <https://www.nhs.uk/apps-library/>
- doi: 10.21037/mhealth.2019.08.05
- Cite this article as:** Deave T, Ginja S, Goodenough T, Bailey E, Piwek L, Coad J, Day C, Nightingale S, Kendall S, Lingam R. The Bumps and BaBies Longitudinal Study (BaBBLLeS): a multi-site cohort study of first-time mothers to evaluate the effectiveness of the Baby Buddy app. *mHealth* 2019;5:42.